



King's Research Portal

DOI:

[10.14301/lcs.v8i4.448](https://doi.org/10.14301/lcs.v8i4.448)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Vineis, P., Avendano-Pabon, M., Barros, H., Chadeau-Hyam, M., Costa, G., Dijmarescu, M., Delpierre, C., D'Errico, A., Fraga, S., Giles, G., Goldberg, M., Zins, M., Kelly-Irving, M., Kivimaki, M., Lang, T., Layte, R., Mackenbach, J. P., Marmot, M., McCrory, C., ... Vollenweider, P. (2017). The biology of inequalities in health: The LIFEPAH project. *Life Course and Longitudinal Studies*, 8(4), 417-439.
<https://doi.org/10.14301/lcs.v8i4.448>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

DOI: <http://dx.doi.org/10.14301/llcs.v8i4.448>

Publisher's PDF, also known as Version of record

Citation for published version (APA):

Paolo Vineis, Mauricio Avendano-Pabon, Henrique Barros, Marc Chadeau-Hyam, Giuseppe Costa, Michaela Dijmarescu, Cyrille Delpierre, Angelo D'Errico, Silvia Fraga, Graham Giles, Marcel Goldberg, Marie Zins, Michelle Kelly-Irving, Mika Kivimaki, Thierry Lang, Richard Layte, Johan P. Mackenbach, Michael Marmot, Cathal McCrory, Cristian Carmeli, Roger L. Milne, Peter Muennig, Wilma Nusselder, Silvia Polidoro, Fulvio Ricceri, Oliver Robinson, Silvia Stringhini (2017). **The biology of inequalities in health: the LIFEPAATH project. Longitudinal and Life Course Studies, Vol 8, No 4.**

Paolo Vineis^{1*}, Mauricio Avendano-Pabon², Henrique Barros³, Marc Chadeau-Hyam¹, Giuseppe Costa⁴, Michaela Dijmarescu¹, Cyrille Delpierre⁵, Angelo D'Errico⁶, Silvia Fraga³, Graham Giles⁷, Marcel Goldberg⁸, Marie Zins⁸, Michelle Kelly-Irving⁵, Mika Kivimaki⁹, Thierry Lang⁸, Richard Layte¹⁰, Johan P. Mackenbach¹¹, Michael Marmot⁹, Cathal McCrory¹⁰, Cristian Carmeli¹², Roger L. Milne⁷, Peter Muennig¹³, Wilma Nusselder¹¹, Silvia Polidoro¹⁴, Fulvio Ricceri⁶, Oliver Robinson¹, Silvia Stringhini¹² & The LIFEPAth Consortium (a)

The biology of inequalities in health: the LIFEPAth project

¹Imperial College London, Department of Epidemiology and Biostatistics, London, UK

² King's College London, Department of Social Sciences, Health and Medicine, London, UK

³ EPIUnit – Institute of Public Health University of Porto, Porto, Portugal

⁴ Turin University Medical School, Department of Clinical Science & Biology, Turin, Italy

⁵ UMR 1027 INSERM – University Paul Sabatier, Department of Epidemiology and Public Health Analyses, Toulouse, France

⁶ Department of Epidemiology, ASL TO3, Piedmont Region, Italy

⁷ Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia

⁸ UMS 011 Inserm - UVSQ « Cohortes épidémiologiques en population », 16 avenue Paul Vaillant Couturier, 94 807, Villejuif, France.

⁹ University College London, Department of Epidemiology and Public Health, London, UK

¹⁰ The Irish Longitudinal Study on Ageing (TILDA), Trinity College Dublin, Dublin, Ireland

¹¹ Erasmus MC, University Medical Center Rotterdam, Department of Public Health, Rotterdam, Netherlands.

¹² Institute of Social and Preventative Medicine, Lausanne University Hospital, Switzerland

¹³ Columbia University, Mailman School of Public Health, New York, USA

¹⁴ Italian Institute for Genomic Medicine (IIGM), formerly known as Human Genetics Foundation (HuGeF), Turin, Italy.

*Correspondence to Paolo Vineis, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, Norfolk Place W112AG London, UK - e-mail: p.vineis@imperial.ac.uk

(a) The LIFEPAATH Consortium (in alphabetical order): Harri Alenius, Mauricio Avendano, Henrique Barros, Murielle Bochud, Cristian Carmeli, Luca Carra, Raphaele Castagne, Marc Chadeau-Hyam, Francoise Clavel-Chapelon, Giuseppe Costa, Emilie Courtin, Michaela Dijmarescu, Cyrille Delpierre, Angelo D’Errico, Pierre-Antoine Dugue, Paul Elliott, Silvia Fraga, Martina Gandini, Valerie Gares, Graham Giles, Marcel Goldberg, Dario Greco, Allison Hodge, Michelle Kelly-Irving, Maryam Karimi, Piia Karisola, Mika Kivimaki, Vittorio Krogh, Thierry Lang, Richard Layte, Benoit Lepage, Johan Mackenbach, Michael Marmot, Cathal McCrory, Roger L. Milne, Peter Muennig, Wilma Nusselder, Salvatore Panico, Dusan Petrovic, Silvia Polidoro, Martin Preisig, Olli Raitakari, Ana Isabel Ribeiro, Fulvio Ricceri, Oliver Robinson, Jose Rubio Valverde, Carlotta Sacerdote, Roberto Satolli, Gianluca Severi, Terrence Simmons, Silvia Stringhini, Rosario Tumino, Anne-Clare Vergnaud, Paolo Vineis, Petter Vollenweider, Marie Zins.

Conflict of interest: None declared

Abstract (now 182 words)

Physical health deteriorates more rapidly with age among men and women with lower socioeconomic status (SES) than among those with higher SES. The biological processes underlying these differences are best understood by adopting a life-course approach. In this paper we introduce the European LIFEPATH project designed to investigate SES and ageing as a phenomenon with two broad stages across life: build-up and decline. LIFEPATH adopts a study design that integrates social science and public health approaches with biology (including molecular epidemiology), using (a) national census-based follow-up data to obtain mortality by socio-economic status, (b) 17 well-characterised population cohorts, and (c) omics measurements (particularly epigenomics). The majority of cohorts include “hard” outcomes (diabetes, cancer, CVD, total mortality), and some extensively phenotyped cohorts also include measurements of the functional components of healthy ageing, including frailty, impaired vision, cognitive function, renal and brain function, osteoporosis, sleep disturbances and mental health. All age groups are represented with two birth cohorts, one cohort of adolescents and several cohorts encompassing young adults (age 18 and above). Furthermore, there is a strong representation of elderly subjects in 7 cohorts.

Key words: social inequalities, socioeconomic status, healthy ageing, life-course, omics, biology.

Introduction

The striking difference in healthy ageing, quality of life and life expectancy observed between individuals from different socioeconomic groups is a major societal challenge that Europe is currently facing. Healthy ageing is an achievable goal in society as it is already experienced by individuals in the highest socioeconomic groups. Individuals with high socioeconomic status (SES) experience much better health and healthier ageing than groups with low SES. Healthy ageing is also strongly related to cumulative exposure to harmful events during the life-course, and particularly in sensitive or critical periods. There is strong evidence that (a) the risk of disease is influenced by early exposures, including *in utero* exposures; (b) the life-course is characterised by *critical* periods (during which changes in exposure have long term effects on disease risks) and *sensitive* periods (during which an exposure has stronger effects on development and, hence, disease risk than at other times). The LIFEPAATH project is an example of the integration of social and biological sciences to provide evidence for public health policies directed at meeting the challenges that the socioeconomic gradient and other factors in healthy ageing represent. The present paper describes the design and some initial results.

Context and rationale

LIFEPAATH is based on a few assumptions that we describe below and have been addressed through the design choices we have made.

1. The healthy ageing model

Socioeconomic differences in health are striking and are seen in all European countries (Mackenbach et al., 2008). Physical health deteriorates more rapidly with age for men and women from lower socioeconomic groups than among those from higher socioeconomic backgrounds. In addition, there is a large amount of evidence that this holds true also for mental health (see for instance Melchior et al., 2013).

This creates a progressively increasing health differential between social groups such that the average physical health of a 70-year-old man or woman with high SES is similar to that of a person with low SES (occupational grade) who is eight years younger (Chandola, T., Ferrie, J.,

Sacker, A. and Marmot, M., 2007). These differences, and the biological processes underlying them, cannot be understood without adopting a life-course approach. In the project we describe below we use the revised Strachan-Sheikh Ben-Shlomo, Y. and Kuh, D., 2002) model of life-course functioning (Blane et al, 2013; Kuh, D., 2007, 2014;), to describe ageing over a lifetime. This model presents ageing as a phenomenon with two broad stages: build-up and decline. The 'build-up' stage, from conception and early intra-uterine life to late adolescence or early twenties, is characterised by rapid successions of developmentally and socially sensitive periods. This stage strongly determines subsequent ageing trajectories as it influences the maximum attained level of health. The second stage, starting in early adulthood, is a period of 'decline' from maximum attained health to loss of function, overt disease and death. Lifestyle and exposures during the second stage can influence the rate at which functioning is lost (**Figure 1**- dashed line). We are aware that this working scheme is a simplification and does not incorporate the complex relationships that underlie ageing processes, but we use it as a conceptual tool to organize our life-course observations. (Figure 1 here)

2. The complex causal network underlying SES and health

There is much evidence that low SES across the life-course is associated with poor ageing trajectories and early death (Chetty, R. et al, 2016; House, J.S. et al, 2005; Mackenbach, J. et al 2008; Stringhini, S. et al, 2010). There is evidence that the impact of SES on ageing is partly mediated by risk factors for non-communicable diseases (NCD), but this is not the only pathway. Also, the features of low SES in high-income countries have dramatically changed over the last century. In Europe, low SES is often associated with food abundance (but of poor quality), lack of physical exercise, and psychosocial stress (including sleep deprivation and mental problems) (Dugravot, A. et al, 2010; Kivimäki, M. et al, 2002, 2012, 2015; Stringhini, S. et al, 2010, 2012;). This adds to the traditional stressors, such as physical fatigue related to occupation and housework. Analyses in the Gazel cohort, which is part of Lifepath (Kivimäki, M. et al, 2008; Platts et al, 2016) show that demanding jobs were associated with fewer life years spent in good health.

SES is linked to many determinants of health, including: 1) access to and use of medical care; 2) access to health information; 3) patterns of unhealthy behaviors (smoking, heavy drinking, unhealthy diet, physical inactivity, drug use); 4) exposure to environmental and occupational hazards; 5) exposure to stressful life events; 6) access to resources mediating the physiological consequences of stress (social relationships and support and cultural capital); 7) early life adverse experiences; 8) age and time-related susceptibilities (Abel, T., 2008; James WP et al. 1997; Kelly-Irving, M. et al. 2011; Lantz, P.M. et al. 2005; Lynch, J.W. et al. 1997; Mitchell, R. et al. 2002; Siegrist, J. et al. 2004; van Doorslaer, E. et al. 2006;). The Whitehall studies (part of LIFEPAH) have shown that the SES gradient in unhealthy ageing and cardiovascular disease (CVD) in particular is not completely explained by traditional risk factors, even when they are reduced to the lowest levels for the whole population (Kivimäki, M. et al, 2008; Marmot, M. et al, 2008). In addition, global changes in the economy are accompanied by rising social inequalities. Instability of markets and lack of growth may lead to instability of health indicators and of ageing trajectories as well. The issue of social inequalities (and their health impact) is increasingly being recognized as a global challenge and priority by the global scientific and policy community, as demonstrated by the publication of an issue of the *Science* magazine devoted to socio-economic differentials, their projections, and their impact on health (Chin, G. and Culotta, E., 2014; Piketty, T. and Saez, E., 2014).

3. Role of behavioural risk factors

Though epidemiology has extensively investigated chemical, physical and behavioural risk factors (such as smoking, diet, alcohol, physical exercise, and occupational factors), these still explain only a fraction of the burden of non-communicable diseases (NCD's) and of common disabilities. For example, a large fraction of breast, colon and prostate cancers is not explained by known risk factors, and the causes of common conditions such as cognitive impairment are largely unknown. The United Nations 25x25 strategy for NCDs (aiming to reduce mortality from NCDs by 25% by the year 2025) lists diabetes/obesity, alcohol, physical exercise, tobacco, raised blood pressure and salt as the main targets of preventive programmes. The strategy is limited by (a) referring exclusively to known risk factors, (b) excluding highly prevalent conditions that limit individual functioning, such as cognitive decline and musculoskeletal disorders, and (c) not accounting for the social structure that underpins the distribution of known risk factors and the effectiveness of preventive strategies and policies.

A considerable amount of evidence shows that socioeconomic factors are as important as lifestyle-related risk factors, and chemical and physical agents, in determining healthy ageing (House, J.S. et al, 2005; Kelly-Irving, M. et al, 2013, Stringhini, S. et al, 2010, 2017;). Low SES is one of the strongest predictors of healthy ageing in Europe, where, on average, disability-free life expectancy at age 65 is 4.5 years shorter in the lowest SES group than in the highest group (Majer, I.M., Nusselder, W.J., Mackenbach, J.P. and Kunst, A.E., 2011). Differences in the actual life expectancy to age 65 in the US are even larger because social conditions predict a higher likelihood of premature death due to violence or unintentional injury (Muennig, P., Fiscella, K., Tancredi, D. and Franks, P, 2010).

4. Biological pathways linking determinants to healthy ageing

What is missing in linking overarching determinants of SES with health and poor ageing is an understanding of the intermediate mechanisms and pathways that relate low SES with deterioration of organic parameters. For example, research on immune markers in the Whitehall II study [Appendix 1] has shown that glucocorticoids (a class of steroid hormones) and inflammation may in part explain how the body mediates the effects of low SES thus leading to disease, and this is partly independent of common known risk factors (Stringhini, S. et al, 2013). More generally, recent studies have shown that SES can influence the global physiological dysregulation across the life-course, measured using allostatic load, a measure of biological multisystem wastage (Barboza Solís, C. et al. 2015; Delpierre et al, 2016 Merkin, S.S. et al. 2014; Nicod, E. et al. 2014).

A research approach based on intermediate biomarkers is more powerful in identifying the links between SES and healthy ageing than one based on traditional risk factors alone because (a) it may explain subtle chronic effects acting throughout the life-course (like chronic stress) that are not easily captured by questionnaire-based epidemiology; (b) it allows tracing signals that start in early life through to health effects in later life; (c) it provides an approach to the discovery of new pathways and causes of disease, particularly through the new omic technologies (Box 1). By analysing intermediate biomarkers potentially involved in various diseases, this approach is likely to reveal some common (currently unknown) roots of many NCDs (multi-morbidity), thus improving our ability to implement successful interventions, with a wide range of actions.

Research strategy

1. Aims of the project

The specific objectives of LIFEPATH are:

- To demonstrate that healthy ageing is highly variable in society, due to multiple environmental, behavioural and social circumstances that affect individual life trajectories.
- To improve the understanding of the mechanisms through which healthy ageing pathways diverge by social circumstances, by investigating life-course biological pathways using omic technologies.
- To provide evidence on the reversibility of the poorer ageing trajectories experienced by individuals exposed to the strongest adversities, by using an experimental approach (randomised experiments of an earned income tax credit programme in New York and Atlanta and a "conditional cash transfer" experiment for poverty reduction in New York City carried out by MDRC, the *Manpower Demonstration Research Corporation*); and to analyse the health consequences of the current economic recession in Europe (i.e. changes in social and economic circumstances).
- To provide updated, relevant and innovative evidence to inform future policies.

The project objectives will be accomplished by using different data sources:

- Europe-wide and national surveys (updated to 2014), including EU-27;
- Longitudinal cohorts (across Europe) with intense phenotyping and repeat biological samples;
- Other large cohorts with biological samples
- A large registry dataset with over a million individuals and very detailed information on work trajectories and health;
- A randomized experiment on conditional cash transfer for poverty reduction in New York City.
- A randomised experiment of an earned income tax credit programme in New York and Atlanta

- Publicly available cohort data.

The geographic location of the studies is shown in Figure 2. A detailed description of the cohorts and corresponding datasets is presented in Table 1 (Appendix 1) (not including Europe-wide national surveys). Data are harmonized and integrated to conceptualize healthy ageing as a composite outcome at different stages of life, resulting from life-course environmental, behavioural and social determinants.

(Figure 2 here)

2. Cohorts participating in the study

LIFEPATH includes information and biological samples from 8 longitudinal population-based cohorts with extensive phenotyping and repeat biological samples, 9 large longitudinal population-based cohorts with biological samples, and a very large cohort without biological samples (WHIP, Work and Health Histories Italian Panel, from Italy). The countries that are covered by the cohorts are France, Italy, Portugal, Ireland, UK, Finland, Switzerland and Australia. These cohorts are only a small proportion of all cohorts available in Europe, but we have chosen them for the combination of good measures of socio-economic status, risk factors for NCD and biomarkers already measured (or availability of blood samples for further testing). The majority of cohorts include “hard” outcomes (diabetes, cancer, cardiovascular disease - CVD, total mortality), and the extensively phenotyped cohorts also include several measurements of the functional components of healthy ageing, including frailty, impaired vision, cognitive function, renal and brain function, osteoporosis, sleep disturbances and mental health. All age groups are represented with two birth cohorts, one cohort of adolescents and several cohorts encompassing young adults (age 18 and above). Furthermore, there is a strong representation of elderly subjects in 7 cohorts (Appendix 1. Table 1). This age structure allows us to address the *first challenge*, the lifecourse perspective (model described in Figure 1).

3. Experimental studies to test policy strategies

Data from cohorts are complemented by data from the evaluation of two randomised experiments in the United States carried out by MDRC (Non-profit research organisation

working to improve programs and policies affecting people with low SES) in New York: (a) the *Opportunity NYC–Family Rewards study*, a randomized experiment of conditional cash transfers (CCT) to help families break the cycle of poverty, and the first CCT program in a high-income country. The program was evaluated through a randomized controlled trial involving approximately 4,800 families and 11,000 children, half of whom could receive the cash rewards if they met the required conditions, and half who were assigned to a control group that could not receive the rewards. Data include extensive measures of social and economic conditions as well as extensive health assessments, which the present project plans to complement with additional biological measures. (b) The *Paycheck Plus*, a randomized experiment of tax credits to help adults without dependent children break the cycle of poverty. The Earned Income Tax Credit is a federal program in which low-income individuals and families receive an annual credit that varies according to their working income. The program was evaluated through a randomized controlled trial involving approximately 10,000 participants, with control participants assigned to the current maximum \$500 annual return and experimental participants assigned to a maximum \$2000 return. Data include extensive measures of social and economic conditions and we will collect data on serum cholesterol, C-reactive protein (a marker of inflammation), glycosylated haemoglobin (a marker of diabetes), blood pressure, abdominal circumference, height, and weight after the intervention. These experimental studies are relevant to the *second challenge*, the complexity of causal relationships and the demonstration of reversibility of some effects.

4. Outcomes and definition of healthy ageing

Indicators of healthy ageing are being developed using the wealth of individual information collected in the cohorts participating in LIFEPAATH (Appendix 1. Table 1) as well as from publicly available datasets. Four “hard” indicators of healthy ageing are being used (incident diabetes, cancer, cardiovascular disease, and death), and several functional and physiological indicators are also available (blood pressure, anthropometry, frailty, physical strength, walking speed and grip strength, eye macular health, sleep, cognitive function including dementia, mental health, bone health, renal function, urinary electrolytes and cardiometabolic measures). The World Health Organisation increasingly emphasizes the need to shift from a definition of health based on hard indicators like mortality and diseases, to ability and disability indicators. In addition to being analysed separately, in LIFEPAATH these measures will be incorporated

into composite indicators of (un)healthy ageing. The contribution of each outcome to the composite indicators will be weighted based on the severity of the outcome (with premature death having the highest weight). To achieve this, we will incorporate methods and estimates from other projects such as the Global Burden of Diseases Collaboration (Lim, S.S. et al, 2012), taking into account life stages and gender. Because the GBD system is controversial, we will also test other systems of weighting such as those based on mobility, impairment of usual activities, pain/discomfort and anxiety/depression derived from a systematic review of the literature. Early-life health events (such as sleep disturbances and changes in blood pressure) will be separated from typically late-life problems like frailty, osteoporosis, macular and cognitive impairment.

In LIFEPAATH we have operationally defined healthy ageing with two components. First, through a general statement about healthy ageing and second, through recommendations about variable selection and choice of analyses across the life-course using longitudinal datasets; these are provided below:

- Statement on Healthy Ageing: “Healthy ageing is the optimal state of performance and wellbeing capable for each particular phase of the lifecourse that can be expected in a society, across social and cultural groups of a population”.
- Key to variable selection is that the component variables be: a) appropriate to the life-course stage, and b) the wellbeing variables capture something of perception and lived experience.

5. Risk factors and healthy ageing

One of the leading hypotheses of LIFEPAATH is that SES operates partially through an unequal distribution of conventional risk factors for poor health across SES strata, but that there is also an additional effect of SES on healthy ageing that is not explained by these conventional risk factors (*third challenge*). Therefore, having good information on a number of risk factors for disease is key to ensuring the success of the project. Also crucial, is the ability to encompass the whole trajectory that links SES with healthy ageing by modelling risk factors (and omics, see below) with proper analytic tools and study designs. These include mediation analysis, structural equation models, randomized experiments and natural experiments. Once the data are harmonized across cohorts, LIFEPAATH will assess:

- (a) the relationship between SES measures and risk factors;
- (b) the relationship between individual risk factors for poor health and their combinations with multiple diseases and (composite) indicators of healthy ageing;
- (c) the intermediate role of risk factors in the relationship of SES with healthy ageing (including interaction between risk factors);
- (d) the proportion of the variance in healthy ageing that remains unexplained by measured risk factors.

6. Biological embedding and omic biomarkers over the life span

Differentiating the analysis of ageing in terms of the two stages of build-up and decline allows an appropriate and detailed exploration of ageing mechanisms, and the use of iterative biological measures from cohorts at different ages. For example, an analysis of the first 'build-up' stage may use as its embodiment measure a biological indicator collected in childhood or early adulthood; whereas an analysis of the 'decline' stage may treat mid-life biological indicators as mediating variables, and measures of physical functioning in later life or diseases as an outcome. This dynamic approach to life-course analyses is central in LIFEPAATH and will be made possible by the integrated use of longitudinal cohorts capturing different life stages with repeat measures and biological samples.

(Box 1 here)

Box 1. What are omics technologies?

Omic technologies allow us to measure the whole set of compounds in a certain compartment (e.g. proteomics is about all proteins). These are the definitions of the omics available or measured *de novo* in LIFEPAATH:

Epigenomics: The analysis of epigenetic changes in DNA, histones, and chromatin that regulate gene expression. Epigenetic changes are changes other than changes in DNA sequence that are involved in gene silencing.

Metabolomics: The scientific study of small molecules (metabolites) that are created from chemicals that originate inside the body (endogenously) or outside the body (exogenously). For purposes of the present report, metabolomics is assumed to include exogenous chemicals found in biological systems in their unmetabolized forms.

Proteomics: The analysis of the proteins produced by cells, tissues, or organisms. Analysis is conducted to understand the location, abundance, and post-translational modification of proteins in a biological sample.

Transcriptomics: Qualitative and quantitative analysis of the transcriptome, that is, the set of transcripts (mRNAs, noncoding RNAs, and miRNAs) that is present in a biological sample.

By combining existing information from birth and adult longitudinal studies, LIFEPAATH is organised into three main interacting streams, illustrated in **Figure 3**:

(a) ***an outer layer of the project: overarching DETERMINANTS***. The focus is on healthy ageing as a continuous process in life ("life-path") and on SES differentials as key determinants, based on the well-documented assumption that SES over time is one of the strongest predictors (and also a summary measure) of the development and preservation of good health and of later-life diseases and disabilities;

(b) ***a middle layer of the project: RISK FACTORS***. The project studies the influences SES has on healthy ageing and life expectancy via behavioural, occupational, nutritional, environmental and other modifiable risk factors for poor health, from childhood to adulthood;

(c) ***an inner layer of the project: BIOLOGICAL PATHWAYS and human social genomics.***

(Figure 3 here)

Omics data (in particular epigenomics, metabolomics and transcriptomics), is already available in several cohorts, and is being used to follow and compare the lifetime ageing pathways of individuals in different SES groups or with different exposure to risk factors from across the life-course. This will allow us to reconstruct biologically-embedded life-course ageing trajectories through the integration of (i) early- and later-life SES and risk factors, and (ii) early- and later-life omic measurements and markers of inflammation or immune response.

The scheme in **Figure 3** explains the different components and outcomes of the project, as well as the types of biomarkers that can be developed as intermediate steps between socio-economic status (SES), risk factors and unhealthy ageing: (a) epigenetic markers, both short-term (and amenable to modification), and long-term (irreversible changes that can be attributed to early-life and late-life exposures); (b) markers of stress and HPA-axis

dysregulation (hypothalamic–pituitary–adrenal axis, a major neuroendocrine system); (c) markers of inflammation and immune response (particularly important in cardiometabolic disease and cancer); (d) markers of neural function and structure.

The following biomarkers are already available or will be measured with existing funds: methylome (n=4,250)(see Box 1); inflammation markers (n=61,000); metabolomics (n=23,000)(Box 1). Most cohorts have C - reactive protein (CRP), and many have cytokines measurements (small proteins involved in cell signaling as immunomodulating agents).

New methylome analyses will be performed for 2,500 new subjects and new transcriptomic analyses for about 600 subjects. Peripheral blood is a convenient source for epigenetic testing, but cellular heterogeneity can confound or mask the results (Adalsteinsson, B.T. et al, 2012; Consortium TEP, 2012; Liu, Y. et al, 2013), because epigenetic signatures differ from one cell subtype to another (Reinius, L.E. et al, 2012). Moreover, the composition of the blood cells can change under a plethora of pathophysiological conditions, and this can also be influenced by SES. The importance of measuring epigenetic patterns across specific cell subtypes also depends on disease status, and subtle differences are likely more important for multifactorial traits. For these reasons, the first analyses of the methylome will be performed in two specific blood cell types in 500 Whitehall II subjects. We will then validate the findings in PBMC (peripheral blood mononuclear cell) in cohorts with cell counts available. Also, the biomarker stock will be enriched by measuring IL6 (interleukin that acts as pro-inflammatory cytokine) and other cytokines in 6,600 additional subjects in Whitehall II (funded by the US National Institutes of Health) and in Skipogh (Table 1), in which also the methylome and RNA sequencing will be measured (n=750).

In some cohorts we also have *genetic data* (Appendix 1. Table 1). This allows three types of analyses: first, we can measure the relative importance of genetic predisposition to disease vs environmental variables including SES; second, we can look at gene-environment interactions with our non-genetic variables including SES (low-penetrance susceptibility); third, we can apply the concept of Mendelian Randomization, i.e. identify gene variants that are associated with intermediate factors that connect SES with health outcomes, and reinforce causal reasoning (given that no genetic variants have been clearly identified that directly influence SES itself). Genotyping will be performed (or is already available) through other funds for 17,000 subjects from Airwave, 6,600 from Whitehall II, 1,100 from Skipogh and 6,000 from Colaus (Table 1).

7. Statistical Methods

This project is working towards developing novel statistical models integrating the different features determining healthy life trajectories. Statistical models are designed to (a) test different conceptual pathways to (un)healthy ageing, integrating information from risk factors, SES, omics and outcomes that contribute to healthy ageing; (b) provide estimates of the weight of different causes, distal/fundamental or proximal, to enable the simulation of different scenarios of public health policies or interventions which might improve healthy living.

(Table 2 here)

An outline of the different aspects of statistical analyses is provided in **Table 2**. Statistical methods are also being developed to investigate the association between SES and healthy ageing, between SES and risk factors, and between risk factors and healthy ageing indices as conceptualised above. Additional methodological efforts will rely on the development of omics profiling techniques integrating data arising from different platforms, and mechanistic models investigating the interplay of these different sets of markers.

Omic signatures - As summarised in Table 3, numerous analyses rely on the identification, from high dimensional omics profiles, of potential omic signatures of SES-related exposures/patterns.

Biological mechanisms – The full exploration of biological mechanisms involved in the SES-affected regulatory cascades relies on in-depth exploration of the correlation structures and the potential effect mediations among the candidate markers found within and across the different platforms.

Time-related patterns - In order to exploit longitudinal data available in LIFEPAH, specific time-related patterns are sought, which are predictive of the quality of ageing.

Conceptual model of ageing and predictive score - At the end of the process, LIFEPAH will construct a holistic conceptual model of ageing by confronting it in an iterative way to data from cohorts involved in the project. This model will integrate the three layers (*outer*:

determinants; *intermediate*: risk factors; *inner*: biomarkers and omics), so that plausible evidence-based pathways encompassing the layers can be built.

8. Age and sex/gender disparities

From birth onwards, a sex-based mortality divide is present, with males having a greater mortality risk than females. Conversely, women have longer life-expectancies but live with higher rates of morbidity and disability. These differences vary in their nature across the lifespan. Sex and gender are central to any research into the processes of ageing, healthy or otherwise. The difficult task in disentangling sex and gender has been highlighted in epidemiology by Krieger (2003) who advocates more complex formulations such as “biologic expressions of gender” and “gendered expression of biology”. Here, we conceptualise sex/gender “as a domain of complex phenomena *that are simultaneously biological and social*, rather than a domain in which the social and biological “overlap”” (Springer, K.W., Hankivsky, O. and Bates, L.M., 2012).

Policy aims

LIFEPATH is devoted to develop policy implications of the scientific findings. Early work in this project will develop conceptual models of the intermediate mechanisms and pathways through which SES gradients influence organic parameters. These models will be empirically examined using data from the study countries to produce evidence of the importance of different pathways in explaining SES inequalities in chronic disease occurrence. In particular, we are interested in the reversibility of certain biological changes (such as inflammation and epigenetics), and in the identification of sensitivity windows in the life course that may inform policies. For example, many policy interventions to mitigate the effects of poverty have so far addressed elderly people, but evidence is emerging that children and even foetuses are susceptible to the multiple effects of low SES (Kramer, M. S., Seguin, L., Lydon, J. and Goulet, L., 2000), and early damage can be carried on to the adult life. The LIFEPATH project takes the innovative step of combining biomarker results with the existing scientific and policy literatures, with the aim of developing a health impact assessment framework (HIA) which will allow the research team to simulate the impact of different policy choices on the overall

occurrence of disease and its distribution across different SES groups. A number of HIA models for specific disease groups are already available. These will be further developed to capture healthy ageing using the results from the LIFEPAATH project. In particular, a model using Markov chains has been developed by the Erasmus group (DynamoOHIA) for application to routine data. The model can be expanded and enriched with socio-economic variables from the LIFEPAATH cohorts and it can be tested among the different European cohorts.

Early achievements so far

We describe here a few early achievements of the project that support some of the basic assumptions and respond to the first challenges we have identified in LIFEPAATH.

1. SES variables and harmonization

The LIFEPAATH consortium brings together data from early life and adult European cohorts with intense phenotyping and repeat biological samples, other large cohorts with biological samples and a large registry dataset with over a million individuals and very rich information on work trajectories and health (D’Errico, A. et al, 2017). To merge and analyse together data from the different LIFEPAATH cohorts, information on occupational class, education, father’s occupational class and income has been harmonized. The harmonization was performed in the initial phases of the project for each of the early life cohorts and adult cohorts participating in the study (Appendix 1. **Table 1**). Detailed codebooks have been created and given that SES indicators relevant for adults differ from those important to characterize children SES, the cohorts have been harmonised using different methods.

Amongst men, the proposed three-level classification of occupational class based on the European Socio-economic Classification (ESeC) – higher, intermediate and lower professions – and education appear not to differ substantially from more detailed classifications in discriminating between main social strata, and in predicting differences in mortality between

them (D'Errico, A. et al., 2017). In spite of differences in recruitment among the cohorts, especially in terms of time, age, gender composition and type of sample, variability in the distribution of the socioeconomic indicators in the different study populations was relatively low. SES indicators categories observed in the different cohorts were also roughly comparable in their distribution to those observed in other studies conducted in Europe, including several conducted on representative samples of the general population. Despite the categories employed, both education and occupational class differed somehow from those used in the present study and the study populations did not share the same characteristics of those included in LIFEPAH. A shift toward higher occupational classes was noted between fathers' and subjects' occupations in most cohorts, which seem to reflect an increasing trend of social improvement in these populations.

Strong differences in mortality between genders were observed for all SES indicators, with much higher and significant associations in males than females, although characterized by variable strength in the different cohorts (**Figure 4**).

(Figure 4 here)

2. SES as a risk factor of physiological wear and tear

Understanding how human environments affect our health by "getting under the skin" and penetrating the cells, organs and physiological systems of our bodies is a key tenet in public health research. In LIFEPAH we examine the idea that early life socio-economic position can be biologically embodied, potentially leading to the production of health inequalities across population groups. Allostatic load (AL), a composite measure of overall physiological wear-and-tear, could allow for a better understanding of the potential biological pathways playing a role in the construction of the social gradient in adult health. We have investigated the factors mediating the link between two components of parental socio-economic position, maternal education (ME) and parental occupation (PO), and AL at 44 years. Data was used from 7,573 members of the 1958 British birth cohort follow-up to age 44. AL was constructed using 14 biomarkers representing four physiological systems. We assessed the contribution of financial/materialist, psychological/psychosocial, educational, and health behaviors/BMI pathways over the life course, in mediating the associations between ME, PO and AL. ME and

PO were mediated by three pathways: educational, material/financial, and health behaviors for both men and women. A better understanding of embodiment processes leading to disease development may contribute to developing adapted public policies aiming to reduce health inequalities (Barboza Solis, C., Fantin, R., Kelly-Irving, M., and Delpierre, C. , 2016). In the same cohort we also showed that AL was associated with subsequent health status (Barboza Solis, C., Fantin, R., Kelly-Irving, M. and Delpierre, C., 2016).

3. SES as a risk factor of high grade inflammation

To explore further potential biological embedding and the consequences of socio-economic position experiences from early life to adulthood, we investigated how socio-economic position indicators at different points across the life course may be related to a combination of 28 inflammation markers. Using blood-derived inflammation profiles measured by a multiplex array in 268 participants from the Italian component of the European Prospective Investigation into Cancer and Nutrition cohort, we evaluated the association between early life, young adulthood and later adulthood socio-economic position with each inflammatory marker separately, or by combining them into an inflammatory score. We identified an increased inflammatory burden in participants whose father had a manual occupation, through increased plasma levels of CSF3 (Granulocyte-colony stimulating factor; $\beta = 0.29$; $P = 0.002$), and an increased inflammatory score ($\beta = 1.96$; $P = 0.029$). Social mobility was subsequently modelled by the interaction between father's occupation and the highest household occupation, revealing a significant difference between "stable Non-manual" profiles over the life course versus "Manual to Non-manual" profiles ($\beta = 2.38$, $P = 0.023$). Low socio-economic position in childhood is associated with modest increase in adult inflammatory burden, which is more evident for stable low SES across life (Castagne, R. et al, 2016).

Outstanding issues

LIFEPATH is a large and ambitious undertaking, and it will try to address some of the main challenges in the field, but it is foreseen that some issues will remain outstanding:

- Measurement error is inherent in SES but also in biomarkers and omics, and this is likely to lead to blurring of the causal pathways
- In spite of the development of statistical methods for mediation analysis, the complex (multi-layered) nature of the relationships between SES, risk factors, biomarkers and outcomes makes it unlikely that a complete picture can be developed. This includes the heterogeneity of the cohorts involved.
- How far our inferences can be extended to other populations with a different distribution of risk factors, SES composition and life expectancy is unknown.

Access to LIFEPATH resources

LIFEPATH aims to become an open-access source of data for biosocial research. The central database contains harmonized variables for socio-economic status, risk factors, biomarkers and outcomes. Access will be regulated by the Steering Committee and rules for access will be made available to interested researchers via the Principal Investigator.

Conclusions: the marriage between biology and social sciences

Accelerated ageing and many human diseases result from a complex interaction between social and biological factors. Risky behaviours, occupational and environmental exposures and psychosocial stress are rooted in societal structures, have historical foundations, and lead to alterations in physiological states (such as allostatic load) that are a prelude to overt disease. The study of the interplay of social and biological factors in LIFEPATH, poses several challenges:

Conceptual challenges: the definition of socio-economic differentials depends on different theoretical constructs that put variable emphasis on status, job position (social class), prestige, and other factors. Similar challenges are encountered in biological modelling, in relation to the relevance of certain markers (such as DNA methylation or metabolomics) to the exploration of ageing trajectories.

Practical challenges, such as the availability and harmonization of good and well measured descriptors of SES across the cohorts. The same challenges are encountered in biomarker

science, where measurements are affected by “nuisance parameters” (such as batch) and confounders, and also need harmonization.

A major challenge is to avoid reductionism in the interpretation of the findings. We seek to establish causal relations across a spectrum that encompasses societal structures, inter-individual relationships, socio-economic position, exposure to risk factors, and the underlying layers of biological mechanisms (that in turn involve whole individuals, tissues, cells and molecules). Causal models have been proposed separately in social sciences and in biological sciences, and they tend to be very different. Our ambition is to integrate them not only conceptually but also operationally and technically in the use of statistical modelling) (Vineis, P., Illari, P. and Russo, F., 2017).

Acknowledgements

This project has received funding from the *European Union’s Horizon 2020 Research and Innovation Programme* under Grant Agreement No. 633666.

References:

- Abel T.** (2008) Cultural capital and social inequality in health. *J Epidemiol Community Health*; 62:e13. ;
- Adalsteinsson, B. T., Gudnason, H., Aspelund, T., Harris, TB, Launer, LJ., Eiriksdottir, G., Smith, A.V. & Gudnason, V.** (2012) Heterogeneity in white blood cells has potential to confound DNA methylation measurements. *PLoS ONE*;7(10), e46705.
- Alves, L., Azevedo, A, Barros, H., Vollenweider, P., Waeber, G. and Marques-Vidal, P.** (2015) Prevalence and management of cardiovascular risk factors in Portuguese living in Portugal and Portuguese who migrated to Switzerland. *BMC Public Health*, 15:307.
- Barboza Solís, C., Fantin, R., Kelly-Irving, M. & Delpierre. C.** (2016) Physiological wear-and-tear and later subjective health in mid-life: Findings from the 1958 British birth cohort. *Psychoneuroendocrinology*.18;74:24-33.
- Barboza, S., Kelly-Irving, M., Fantin, R., Darnaudery, M., Torrisani, J., Lang, T. & Delpierre, C.** (2015) Adverse childhood experiences and physiological wear-and-tear in midlife: Findings from the 1958 British birth cohort. *Proc Natl Acad Sci USA*. E738-E746;
- Ben-Shlomo., Y. and Kuh, D. (Strachan, D. and Sheikh, A., 2002)** (Eds). *A Life Course Approach to Chronic Disease Epidemiology. 2nd edition (pp. 240-259). Oxford: Oxford University Press.*
- Blane, D., Kelly-Irving, M., d’Errico, A., Bartley, M. & Montgomery, S.** (2013) Social-biological transitions: how does the social become biological? *Longitudinal and Life Course Studies* 4 (2):136 – 146.
- Castagne, R., Delpierre, C., Kelly-Irving, M., Campanella, G., Guida, F., Krogh, V., Palli, D., Panico, S., Sacerdote, C., Tumino, R., Kyrtopoulos, S., Hosnijeh, F.S., Vermeulen, R., Vineis, P., Stringhini, S. and Chadeau-Hyam, M.** (2016) A life course approach to explore the biological embedding of socioeconomic position and social mobility through circulating inflammatory markers. *Sci Rep*. 6:25170
- Chandola, T., Ferrie, J., Sacker, A. & Marmot, M.** (2007) Social inequalities in self-reported health in early old age: follow-up of prospective cohort study. *BMJ*. 334(7601):990;
- Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Berengon, A. & Cutler, D.** (2016) The association between income and life expectancy in the United States, 2001-2014. *315*(16):1750-66.
- Chin, G. & Culotta, E.** (2014) The science of inequality. What the numbers tell us. Introduction. *Science*. 344(6186):818-21.
- Consortium TEP.** (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*; 489(7414), 57–74.

Correia, S., Rodrigues, T. and Barros, H. (2014) Socio-economic variations in female fertility impairment: a study in a cohort of Portuguese mothers. *BMJ*; 4(1):e003985.

Cronin, H., O'Reagan, C., Finucane, C., Kearney, P. and Kenny, R.A. (2013) Health and aging: development of the Irish Longitudinal Study on Ageing health assessment. *J Am Geriatr Soc*; 61 Suppl 2:S269-78.

D'Errico, A., Ricceri, F, Stringhini, S., Carmeli, C., Kivimaki, M., Bartley, M., McCrory, C., Bochud, M., Vollenweider, P., Tumino, R., Goldberg, M., Zins, M., Barros, H., Giles, G., Severi, G., Costa, G., Vineis, P. and the LIFEPAATH Consortium. (2017) Socioeconomic indicators in epidemiologic research: A practical example from the LIFEPAATH study. *PLoS One*; 12 (5):e0178071.

Dartois, L., Fagherazzi, G., Boutron-Ruault, M.C., Mesrine, S. and Clavel-Chapelon, F. (2014) Association between five lifestyle habits and cancer risk: results from the E3N cohort. *Cancer Prev Res (Phila)*. 7(5):516-25.

Delpierre C. (2016) Origins of health inequalities: the case of allostatic load. *Longitudinal and Life Course Studies*; 7, 79-102.

Dugravot, A., Sabia, S., Stringhini, S., Kivimaki, M., Westerlund, H., Vahtera, J., Gueguen, A., Zins, M., Goldberg, M., Nabi, H. & Singh-Manoux, A. (2010) Do socioeconomic factors shape weight and obesity trajectories over the transition from midlife to old age? Results from the French GAZEL cohort study. *Am J ClinNutr*;92(1):16-23.

Firmann, M., Mayor, V., Marques-Vidal, P., Bochud, M., Pecoud, A., Hayoz, D., Paccaud, F., Preisig, M., Song, K.S., Yuan, X., Danoff, T.M., Stirnadel, H.A., Waterworth, D., Mooser, V., Waeber, G. and Vollenweider. P. (2008). The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovascular Disorders*. 8:6.

Gallo, V., Mackenbach, J.P., Ezzati, M., Menvielle, G., Kunst, A.E., Rohrmann, S., Kaaks, R., Teucher, B., Boeing, H., Bergmann, M.M., Tionneland, A., Dalton, S.O., Overvad, K., Redondo. M.L., Aguado, A., Daponte, A., Arriola, L., Navarro, C., Gurrea, A.B., Khaw, K.T., Wareham, N., Ket, T., NAska, A., Tricopoulou, A., Trichopoulos, D., Masala, G., Panico, S., Contiero, P., Tumino, R., Bueno-de-Mesquita, H.B., Siersema, P.D., Peeters, P.P., Zackrisson, S., Almgvist, M., Eriksson, S., Hallmans, G., Skeie, G., Braaten, T., Lund, E., Illner, A.K., Mouw, T., Riboli, E. & Vineis, P. (2012) Social inequalities and mortality in Europe – results from a large multi-national cohort. *PLoS One*.7(7):e39013.

Hodge, A.M., Flicker, L., O'Dea, K., English, D.R. and Giles, G.G. (2013) Diabetes and ageing in the Melbourne Collaborative Cohort Study (MCCS). *Diabetes Res Clin Pract*. 337(8754):1387-93.

House, J.S., Lantz, P.M. & Herd, P. (2005) Continuity and change in the Social Stratification in Aging and Health over the Life Course: Evidence from a Longitudinal Study from 1986 to

2001/2002 (Americans' Changing Lives Study). *J Gerontol B Psychol Sci Soc Sci.* 60 Spec No 2:15-26;

Hyppönen E, Davey Smith G, Shepherd P, Power C. (2005) An intergenerational and lifecourse study of health and mortality risk in parents of the 1958 birth cohort: (II) mortality rates and study representativeness. *Public Health* 119(7):608-15

James, W.P., Nelson, M., Ralph, A. & Leather, S. (1997) Socioeconomic determinants of health. The contribution of nutrition to inequalities in health. *Br Med J* ;314:1545-9;

Kearney, P.M., Cronin, H., O'Reagan, C., Kamiya, Y., Savva, G.M., Whelan, B. and Kenny, R. Cohort profile: the Irish Longitudinal Study on Ageing. *Int J Epidemiol* 2011. 40(4):877-884.

Kelly-Irving, M., Delpierre, C., Schieber, A.C., Lepage, B., Rolland, C., Afrite, A., Pascal, J., Cases, C., Lombrail, P. & Lang, T. (2011) Do general practitioners overestimate the health of their patients with lower education? *Soc Sci Med* ;73:1416-21;

Kelly-Irving, M., Lepage, B., Dedieu, D., Bartley, M., Blane, D., Grosclaude, P., Lang, T. & Delpierre, C. (2013) Adverse childhood experiences and premature all-cause mortality. *Eur J Epidemiol.* 28(9):721-34.

Kivimäki, M., Head, J., Ferrie, J.E., Singh-Manoux, A., Westerlund, H., Vahtera, J., Leclerc, A., Melchior, M., Chevalier, A., Alexanderson, K., Zins, M. & Goldberg, M. et al. (2008) Sickness absence as a prognostic marker for common chronic conditions: analysis of mortality in the GAZEL study. *Occup. Environ. Med.*; 65:820-826.

Kivimäki, M., Jokela, M., Nyberg, S.T., Singh-Manoux, A., Fransson, E.I., Alfredsson, L., Bjorner, J., Borritz, M., Burr, H., Casini, A., Clays, E., De Bacquer, D., Dragano, N., Erbel, R., Geuskens, G.A., Hamer, M., Hoofman, W.E., Houtman, I.L., Jockel, K.H., Kittel, F., Knutsson, A., Koskenvuo, M., Lunau, T., Madsen, E.E.H., Nielsen, M.L., Nordin, M., Oksanen, T., Pejtersen, J.H., Pentti, J., Rugulies, R., Salo, P., Shipley, M.J., Siegrist, J., Steptoe, A., Suominen, S.B., Theorell, T., Vahtera, J., Westerholm, P.J.M., Westerlund, H., O'Reilly, D., Kumari, M., Batty, D.G., Ferrie, J.E. & Virtanen, M. (2015) Long working hours and risk of coronary heart disease and stroke: a systematic review and meta-analysis of published and unpublished data for 603 838 individuals. *Lancet* ;386:1739-46).

Kivimäki, M., Leino-Ajas, P., Luukkonen, R., Riihimäki, H., Vahtera, J. & Kirjonen, J. (2002) Work stress and risk of cardiovascular mortality: prospective cohort study of industrial employees. *BMJ.* 325(7369):857;

Kivimäki, M., Shipley, M.J., Ferrie, J.E., Singh-Manoux, A., Batty, G.D., Chandola, T., Marmot, M.G. & Davey Smith, G. (2008) Best-practice interventions to reduce socioeconomic inequalities of coronary heart disease mortality in UK: a prospective occupational cohort study. *Lancet*;372(9650):1648-54.

Kramer, M. S., Seguin, L., Lydon, J., & Goulet, L. (2000). Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly? *Paediatric and perinatal epidemiology*, 14(3), 194-210.

Krieger, N., Chen, J.T., Waterman, P.D., Rehkopf, D.H. & Subramanian, S.V. (2003) Race/ ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures – the public health disparities geocoding project. *Am J Public Health*. 93(10):1655-71;

Kuh, D. (2007) A life course approach to healthy ageing, frailty, and capability. *J Gerontol A Biol Sci Med Sci*. 62:717-21.

Lantz, P.M., House, J.S., Mero, R.P. & Williams, D.R. (2005) Stress, life events, and socioeconomic disparities in health: results from the American's Changing Lives Study. *Journal of Health & Social Behavior* ;46:274-88.

Larsen, P. S., Kamper-Jorgensen, M., Adamson, A., Barros, H., Bonde, J. P., Brescianini, S., Brophy, S., Casas, M., Charles, M.A., Devereux, G., Eggesbo, M., Fantini, M.P., Frey, U., Gehring, U., Grazuleviciene, R., Henriksen, T.B., Hertz-Picciotto, I., Heude, B., Hryhorczuk, D.O., Inskip, H., Jaddoe, V.W., Lawlor, D.A., Ludvigsson, J., Kelleher, C., Kiess, W., Koletzko, B., Kuehni, C.E., Kull, I., Kyhl, H.B., Magnus, P., Momms, I., Murray, D., Pekkanen, K., Porta, D., Poulsen, G., Richiardi, L., Roeleveld, N., Skovgaard, A.M., Sram, R.J., Strandberg-Larsen, K., Thijs, C., Van Eijsden, M., Wright, J., Vrijheid, M. & Andersen, A.M. (2013) Pregnancy and birth cohort resources in europe: a large opportunity for aetiological child health research. *Paediatr Perinat Epidemiol*; 27(4), 393-414.

Layte R, McCrory C. (2013) Paediatric chronic illness and educational failure: the role of emotional and behavioural problems. *Soc Psychiatry Psychiatr Epidemiol*. 48(8):1307-16

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Amann M, Anderson HR, Andrews KG, Aryee M, Atkinson C, Bacchus LJ, Bahalim AN, Balakrishnan K, Balmes J, Barker-Collo S, Baxter A, Bell ML, Blore JD, Blyth F, Bonner C, Borges G, Bourne R, Boussinesq M, Brauer M, Brooks P, Bruce NG, Brunekreef B, Bryan-Hancock C, Bucello C, Buchbinder R, Bull F, Burnett RT, Byers TE, Calabria B, Carapetis J, Carnahan E, Chafe Z, Charlson F, Chen H, Chen JS, Cheng AT, Child JC, Cohen A, Colson KE, Cowie BC, Darby S, Darling S, Davis A, Degenhardt L, Dentener F, Des Jarlais DC, Devries K, Dherani M, Ding EL, Dorsey ER, Driscoll T, Edmond K, Ali SE, Engell RE, Erwin PJ, Fahimi S, Falder G, Farzadfar F, Ferrari A, Finucane MM, Flaxman S, Fowkes FG, Freedman G, Freeman MK, Gakidou E, Ghosh S, Giovannucci E, Gmel G, Graham K, Grainger R, Grant B, Gunnell D, Gutierrez HR, Hall W, Hoek HW, Hogan A, Hosgood HD 3rd, Hoy D, Hu H, Hubbell BJ, Hutchings SJ, Ibeanusi SE, Jacklyn GL, Jasrasaria R, Jonas JB, Kan H, Kanis JA, Kassebaum N, Kawakami N, Khang YH, Khatibzadeh S, Khoo JP, Kok C, Laden F, Lalloo R, Lan Q, Lathlean T, Leasher JL, Leigh J, Li Y, Lin JK, Lipshultz SE, London S, Lozano R, Lu Y, Mak J, Malekzadeh R, Mallinger L, Marcenes W, March L, Marks R, Martin R, McGale P, McGrath J, Mehta S, Mensah GA, Merriman TR, Micha R, Michaud C, Mishra V, Mohd Hanafiah K, Mokdad AA, Morawska L, Mozaffarian D, Murphy T, Naghavi M, Neal B, Nelson PK, Nolla JM, Norman R, Olives C, Omer SB, Orchard J, Osborne R, Ostro B, Page A, Pandey

KD, Parry CD, Passmore E, Patra J, Pearce N, Pelizzari PM, Petzold M, Phillips MR, Pope D, Pope CA 3rd, Powles J, Rao M, Razavi H, Rehfuss EA, Rehm JT, Ritz B, Rivara FP, Roberts T, Robinson C, Rodriguez-Portales JA, Romieu I, Room R, Rosenfeld LC, Roy A, Rushton L, Salomon JA, Sampson U, Sanchez-Riera L, Sanman E, Sapkota A, Seedat S, Shi P, Shield K, Shivakoti R, Singh GM, Sleet DA, Smith E, Smith KR, Stapelberg NJ, Steenland K, Stöckl H, Stovner LJ, Straif K, Straney L, Thurston GD, Tran JH, Van Dingenen R, van Donkelaar A, Veerman JL, Vijayakumar L, Weintraub R, Weissman MM, White RA, Whiteford H, Wiersma ST, Wilkinson JD, Williams HC, Williams W, Wilson N, Woolf AD, Yip P, Zielinski JM, Lopez AD, Murray CJ, Ezzati M, AlMazroa MA, Memish ZA. (2012) A comparative risk assessment of burden disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*; 380:2224-60.

Liu, Y., Aryee, M.J., Padyukov, L., Fallin, M.D., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M., Shchetynsky, K., Scheynius, A., Kere, J., Alfredsson, L., Klareskog, L., Ekstrom, T.J. & Feinberg, A.P. (2013) Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nature Biotechnology*. 31, 142–147.

Lynch J.W, Kaplan G.A.& Salonen J.T. (1997) Why do poor people behave poorly? Variation in adult health behaviours and psychosocial characteristics by stages of the socioeconomic lifecourse. *Soc Sci Med* ; 44:809-819.

Lynch, J.W., Kaplan, G.A. & Shema, S.J. (1997) Cumulative impact of sustained economic hardship on physical, cognitive, psychological, and social functioning. *N Engl J Med* ;337:1889-95.

Mackenbach, J., Stirbu, I., Roskam, A.J.R., Schaap, M.M., Menvielle, G., Leinsalu, M. & Kunst, A.E. (2008) Socioeconomic inequalities in health in 22 European Countries. *N Engl J Med*; 358(23):2468-81.

Majer, I.M., Nusselder, W.J., Mackenbach, J.P. & Kunst, A.E. (2011) Socioeconomic inequalities in life and health expectancies around official retirement age in 10 Western-European countries. *J Epidemiol Community Health*. 65(11):972-9. doi: 10.1136/jech.2010.111492;

Mäki, N.E., Martikainen, P.T., Eikemo, T., Menvielle, G., Lundberg, O. & Mackenbach, J. (2014) The potential for reducing differences in life expectancy between educational groups in five European countries: the effects of obesity, physical inactivity and smoking. *J Epidemiol Community Health*. 68(7):635-40.

Marmot, M.G., Shipley, M.J., Hemingway, H., Head, J. & Brunner E.J. (2008) Biological and behavioural explanations of social inequalities in coronary heart disease: The Whitehall II study. *Diabetologia*. 51:1980-8.

- Melchior, M., Chastang, J.F., Head, J., Goldberg, M., Zins, M., Nabi, H. & Younes, N.** (2013) Socioeconomic position predicts long-term depression trajectory: a 13-year follow-up of the GAZEL Cohort Study. *Mol Psychiatry*. 18:112–121.
- Merkin, S.S., Karlamangla, A., Roux, A.V., Shrager, S. & Seeman, T.E.** (2014) Life course socioeconomic status and longitudinal accumulation of allostatic load in adulthood: multi-ethnic study of atherosclerosis. *Am J Public Health*. 104(4):e48-55.
- Mitchell, R., Blane, D. & Bartley, M.** (2002) Elevated risk of high blood pressure: climate and the inverse housing law. *Int J Epidemiol* 31:831-8;
- Muennig, P., Fiscella, K., Tancredi, D. & Franks, P.** (2010) The relative health burden of selected social and behavioral risk factors in the United States: Implications for Policy. *Am J Public Health*. 100(9): 1758-1764.
- Nicod, E., Stringhini, S., Marques-Vidal, P., Paccaud, F., Waeber, G., Lamiraud, K., Vollenweider, P. and Bochud, M.** (2014) Association of education and receiving social transfers with allostatic load in the Swiss population-based CoLaus study. *Prev med*. 63:63-71.
- Pereira M, Lunet N, Paulo C, Severo M, Azevedo A, Barros H.** (2012) [Incidence of hypertension in a prospective cohort study of adults from Porto, Portugal](#). *BMC Cardiovasc Disord*. 28;12:114.
- Piketty, T. & Saez, E.** (2014) Inequality in the long run. *Science*; 344:838-42.
- Platts, L.G., Head, J., Stenholm, S., Singh Chungkham, H., Goldberg, M. & Zins, M.** (2017) Physical occupational exposures and healthy life expectancy in a French occupational cohort. *Occup Environ Med*. 74(3):176-183.
- Raitakari, O., Juonala, M., Kahonen, M., Taittonen, L., Laitinen, T., Maki-Torkko, N, Jarvisalo, M.J., Uhari, M., Jokinen, E., Ronnemaa, T., Akerblom, H.K. & Viikari, J.S.** (2003) Cardiovascular risk factors in childhood and carotid artery intima-media thickness in adulthood: The Cardiovascular Risk in Young Finns Study. *JAMA*. 5:290(17):2277-83.
- Ramos, E. & Barros, H.** (2007) Family and school determinants of overweight in 13-year-old Portuguese adolescents. *Acta Paediatr*;96(2):281-6;
- Reinius, L. E., Acevedo, N., Joerink, M., Pershagen, G., Dahlen, S.E., Greco, D., Soderhall, C., Scheynius, A. & Kere, J.** (2012) Differential DNA Methylation in Purified Human Blood Cells : implications for cell lineage and studies on disease susceptibility. *PLoS ONE*. 7(7), e41361. Epub 2015 Jul 25.
- Shiue I.** (2014) Social correlates of total cancer in adults and the very old: UK Understanding Society Cohort, 2009-2010. *Public Health Journal*. 128 (4): 376-379.
- Siegrist, J. & Marmot, M.** (2004) Health inequalities and the psychosocial environment – two scientific challenges. *Soc Sci Med* 58:1463-73.

Springer, K.W., Hankivsky, O. & Bates, L.M. (2012) Gender and health: relational, intersectional, and biosocial approaches. *Soc Sci Med.* 74(11):1661-6.

Stringhini, S., Batty, G.D., Bovet, P., Shipley, M.J., Marmot, M.G., Kumari, M., Tabak, A.G. & Kivimaki, M. (2013) Association of lifecourse socioeconomic status and chronic inflammation and type 2 diabetes risk: the Whitehall II prospective cohort study. *PLoS Med.* 10(7):e1001479.

Stringhini, S., Carmeli, C., Jokela, M., Avendano, M., Muennig, P., Guida, F., Ricceri, F., d'Errico, A., Barros, H., Bochud, M., Chadeau-Hyam, M., Clavel-Chapelon, F., Costa, G., Delpierre, C., Fraga, S., Goldberg, M., Giles, G.G., Krogh, V., Kelly-Irving, M., Layte, R., Lasserre, A.M., Marmot, M.G., Preisig, M., Shipley, M.J., Vollenweider, P., Zins, M., Kawachi, I., Steptoe, A., Mackenbach, J.P., Vineis, P., Kivimaki, M. (2017) Socioeconomic status and the 25x25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. *The Lancet* ; 389(10075): 1229-1237.

Stringhini, S., David Batty, G., Bovet, P., Shipley, M.J., Marmot. M.G., Kumari, M., Tabak. A.G. and Kivimaki, M. (2013) Association of lifecourse socioeconomic status with chronic inflammation and Type 2 Diabetes risk: the Whitehall II Prospective Cohort Study. *PLos One.* 10(7):e1001479.

Stringhini, S., Sabia, S., Shipley, M., Brunner, E., Nabi, H., Kivimaki, M. & Singh-Manoux, A. (2010) Association of socioeconomic position with health behaviors and mortality. *2010 JAMA.* 303(12):1159-66.

Stringhini, S., Tabak, A.g., Akbaraly, T.N., Sabia, S., Shipley, M.J., Marmot, M.G., Brunner, E.J., Batty, G.D., Bovet, P. & Kivimaki, M. (2012) Contribution of modifiable risk factors to social inequalities in type 2 diabetes: prospective Whitehall II cohort study. *BMJ* ;345:e5452.

van Doorslaer, E., Masseria, C., Koolman, X. & the OECD Health Equity Research Group. (2006) Inequalities in access to medical care by income in developed countries. *CMAJ* ;174:177-83.

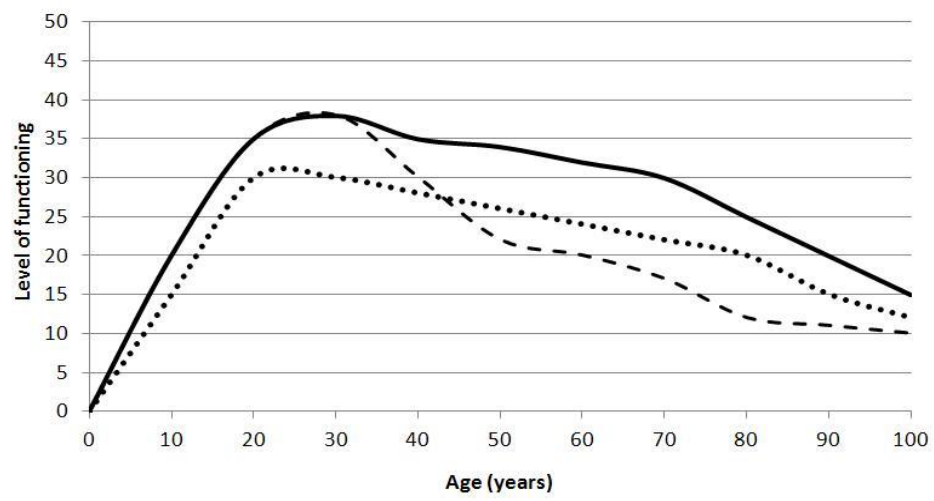
Vineis P, Illari P, & Russo F. (2017) [Causality in cancer research: a journey through models in molecular epidemiology and their philosophical interpretation.](#) *Emerg Themes Epidemiol.* 7;14:7.

Westerlund, H., Kivimaki, M., Singh-Manoux, A., Melchior, M., Ferrie, J.E., Pentti, J., Jokela, M., Leineweber, C., Goldberg, M., Zins, M. and Vahtera, J. (2009) Self-rated health before and after retirement in France (GAZEL): a cohort study. *Lancet.* 374(9705):1889-96.

Whelan, B.J & Savva, G.M. (2013) Design and methodology of the Irish Longitudinal Study on Ageing. *J Am Geriatr.* Suppl. 2:S265-268.

Zins, M., Bonenfant, S., Carton, M., Coeuret-Pellicer, M., Gueguen, A., Gourmelen, J., Nchtigal, M., Ozguler, A., Quesnot, A., Ribet, C., Rodrigues, G., Serrano, A., Sitta, R., Brigand, A., Henny, J. & Goldberg, M. (2010) The CONSTANCES cohort: an open epidemiological laboratory. *BMC Public Health*. 12;10:479.

Figure 1. The “build-up” and “decline” model.



Appendix 1. Table 1 – LIFEPAATH: cohorts included in the consortium. Study design/cohort description

Cohort	Study design	Outcomes and clinical tests	Already available/funded biomarkers
Whitehall II (UK) N=10,308 (Stringhini, S. et al, 2013)	The Whitehall II study was established in 1985 to examine the socioeconomic gradient in health among 10,308 London-based civil servants (6,895 men and 3,413 women) aged 35–55. Baseline examination (phase 1) took place during 1985–1988, and involved a clinical examination and a self-administered questionnaire containing sections on demographic characteristics, health, lifestyle factors, work characteristics, social support, and life events.	Mortality, cancers, CVD events. BP, diabetes, anthropometry, physical functioning, cognitive function, mental health, hospitalizations.	Genotype (metabochip, cardiochip); metabolomics (NMR)(n=6,600); repeat data on lipids, glucose, insulin, inflammatory markers (CRP, IL6), cortisol (saliva, hair) (n=6,000-10,000). Subset: Brain MRI
TILDA (Ireland) N = 8,175 (Kearney, PM et al, 2011); Cronin, H., O'Reagan, C., Finucane, C., Kearney, P. & Kenny, R.A., 2013)	The Irish Longitudinal Study on Ageing (TILDA) is a large prospective cohort study examining the social, economic, and health circumstances of 8,175 community-dwelling older adults (3,744 men, 4,431 women) aged 50 years and older at the time of recruitment and resident in the Republic of Ireland. Baseline examination (phase 1) took place during 2009-2011 and participants are followed up biennially. There were three components to the survey. Respondents completed a computer-assisted personal interview (n = 8,175) and a separate self-completion paper and pencil module (n = 6,915) which collected information that was considered sensitive. All participants were invited to undergo a separate health assessment at one of two national centers using trained nursing staff (n=5,897). A total of 5036 respondents completed the health-centre	BP, heart rate variability, pulse wave velocity, cerebral perfusion, balance, gait speed, osteoporosis, renal function, cognitive function, physical function, accelerometry, muscle mass, muscle strength, retinal imaging, macular pigment density, sensory performance, anthropometry mental health, quality of life, hospitalization, mortality, cancers, diabetes, CVD events, medical history, medications.	Lipids (N=5,800). Funded: Vitamin B12, Vitamin D, CRP, Creatinine, Folate, HbA1c (N=5,800). Subset: Brain MRI

	based assessment and a further 861 respondents completed a home-based assessment which involved a reduced set of tests. A more detailed exposition of study design, sample selection and protocol is available elsewhere (Whelan, B.J. & Savva, G.M., 2013).		
Generation 21 (Portugal) N=8,647 Barros, H., Correia, S. et al. BMJ Open 2014; Larsen PS et al., 2013	Generation 21 (G 21) comprises a cohort of 8,647 newborns recruited in 2005–2006 in the Porto Metropolitan Area, in northern Portugal. Recruitment occurred at the 5 public maternity units, which are responsible for 95% of all births in the region (remaining births occurred at private hospitals) ²¹ . During the hospital stay, women delivering live births were invited to participate, and 92% of mothers agreed. All who agreed were invited to be re-evaluated at child's 4 years of age (2009-2011) , at 7 (2012-2014) and then again at 10 years (2015-2016).	Self-reported health status (health conditions), Tetrapolar Bioimpedance, ECG, Blood Pressure, Spirometry, DEXA, cognitive test.	CRP, lipids, glucose, insulin (4yrs n=1,530; 7yrs n=4,500)
The Airwave Health Monitoring Study (UK) N=45,596 Paul Elliott http://www.police-health.org.uk/	Airwave Health Monitoring Study is an ongoing long term epidemiological occupational cohort study open to all police forces in UK. The study's current main goal is to investigate any possible impacts of TETRA on health by looking at TETRA exposure and subsequent health amongst police officers and staff. The study was launched in 2004, and since its inception, 28 out of 54 police forces have agreed to participate. The cohort aims to have 60,000 participants by 2018. By December 2010, the cohort had recruited 42,057 participants out of which 34,957 have undergone a health screening that includes extensive	Cancer, death, hospitalizations. Sickness absences, blood pressure, ECG, arterial stiffness, anthropometry, cognitive test on sub-samples.	Blood count, haemoglobin, urea, creatinine, Gamma GT, lipids, glucose, HBA1c, C-peptide, CRP, fibrinogen, prothrombin time (N=35,000) Genotyping (N=17,000) Metabolome (N=3,000)

	lifestyle and questionnaire data, clinical measurements and collection of biological samples.		
EPIPORTO (Portugal) N=2,485 (Alves, L., Barros, H., et al., 2015)	The EPIPorto is a general adult population-based cohort established with the initial aim of evaluating the major determinants of cardiovascular health. For this purpose, 2,485 (949 men and 1,536 women) adult dwellers in Porto, aged 18 years or over, were recruited between 1999 and 2003 using random digit dialling. The follow-up studies were performed for the whole sample in 2005-2008, and in 2014-2015. At all waves, information was collected using questionnaires administered by trained interviewers, self-administered questionnaires and objective measurements were made, including physical examination and blood tests (Pereira et al, 2012).	Mortality, cancers, CVD events. BP, diabetes, anthropometry, ECG, Musculoskeletal conditions.	CRP, lipids, glucose, insulin, fibrinogen
SKIPOGH study Switzerland (3 centres) N=1,128 Bochud, M., Ponte, B. et al. Hypertension 2014	The Swiss Kidney Project on Genes in Hypertension is a longitudinal family-based study, following the standardised EPOGH (European Project on Genes in Hypertension) protocol. Baseline examination was conducted between 2009 and 2013. Three-year follow-up examination started in 2013 and is currently ongoing (expected to be finished in 2015). The aim of the study is to explore the role of genes and kidney haemodynamics in blood pressure regulation and kidney function in the general population. From December 2009 to March 2013, adult participants were recruited in two regions (Berne and Geneva) and one city (Lausanne) of Switzerland. A random sample of the inhabitants was drawn using different strategies. Inclusion criteria were (1) having a minimum	Deaths, CVD events, diabetes, hypertension, 24h BP, anthropometry, ECG, arterial properties, retinal imaging, renal function, chronic kidney disease, kidney morphology, cardiac imaging (3D); hand-grip strength.	Methylome (Illumina 450K) (N=250). Transcriptomics (N=250). Inflammation markers, CRP, lipids, fasting glucose, insulin. Urine Na, K, urea, steroid metabolites (N=1,100). Cardiometabochip (200K SNPs) (N=1,100).

	age of 18 years; (2) being of European ancestry; (3) having \geq one and ideally three first-degree family members willing to participate; and (4) providing a written informed consent. Pregnant or breastfeeding women were not included. The general participation rate was 27.1%. At baseline, we collected data on cardiovascular and metabolic risk factors as well as on the prevalence of kidney and cardiovascular diseases. During follow-up, we are collecting data on new kidney and cardiovascular events. The primary endpoints are fatal and non-fatal strokes, ischaemic heart disease, heart failure and chronic kidney disease. The 1,128 participants (537 men and 591 women) belong to 272 nuclear families.		
COLAUS Study N=6,733 (Firmann, M. et al., 2008)	The CoLaus (COhort LAUSannoise) is an ongoing prospective study assessing the clinical, biological and genetic determinants of cardiovascular disease in the city of Lausanne, Switzerland (Firmann et al., 2008). The initial survey was conducted between 2003 and 2006 and included 6,733 participants aged between 35 and 75 years; the first follow-up survey was conducted 5.5 years afterwards and included 5,064 participants. In each survey, data on socio-economic status, lifestyle, mental status and cardiovascular risk factors is collected by questionnaire or clinical examination.	Deaths, CVD events, diabetes, hypertension, BP, anthropometry, physical and cognitive functioning, FFQ, PAFQ. N=2,000 polysomnography (Hypnolaus) N=3,500 psychiatric examination (Psycholaus)	Inflammation markers (N=6,300), CRP, lipids, fasting glucose, insulin, vitamins, genotype (500 K Affymetrix chip technology)
Growing Up in Ireland (Ireland) N = 19,702	Longitudinal child cohort established in 2007/2008. The study takes place over seven years and follows the progress of two groups of children: 1) containing 11,134 participants	Chronic illness of parents and children and whether confirmed by physician.	Measured heights and weights for parents and children at each wave.

Layte, R. and McCrory, C, 2013	aged 9 months at baseline with three study waves (9 months, 3 years and 5 years) and 2) containing 8,568 participants aged 9 years at baseline (9 years, 13 years). The dataset gathered contains multiple SES measures of parents at each wave and economic strain and anthropometric measures at each wave of child and parents.	Acute illness for infants at baseline. Child psychological adjustment, cognitive function, parent/child relationship, child self-concept.	Measured head circumference for 9-month cohort at baseline.
--------------------------------	--	--	---

Other large longitudinal cohorts with biological samples (n=202,555).

Cohort	Study Design	Outcomes and clinical trials	Already available/ funded biomarkers
EPIC- Italy N=47,749 Gallo, V. et al, PLoS One 2012	The European Prospective Investigation into Cancer and Nutrition (EPIC) is a large European study on diet and cancer. The Italian component of EPIC (EPIC-Italy) recruited 47,749 adult volunteers (men and women) at five centres: Varese and Turin in northern Italy, Florence in central Italy and Naples and Ragusa in southern Italy. All participants signed an informed consent form and completed two questionnaires: one about dietary habits (food-frequency) and one about lifestyle, with information on education, socioeconomic status, occupation, history of previous illnesses and surgery, lifetime tobacco use and alcohol consumption and physical activity. EPIC database records were linked to cancer and regional mortality	All cancers, CVD, diabetes, BP.	Methylome (Illumina 450K) for >1,000 subjects. Inflammation markers, CRP, lipids, in subsets

	registries after EPIC database quality control. All EPIC-Italy centres except Naples are covered by population-based cancer registries. In Naples, follow-up information was collected from electronic hospital discharge records and also by periodic personal contact with participants. We are including in LIFEPAATH centres from EPIC Italy (N=34,148) with the exclusion of Florence.		
E3N (France) N=98,995 (Dartois., L., Fagherazzi, G., Boultron-Ruault, M.C., Mesrine, S. and Clavel-Chapelon, F. 2014)	The E3N study (Étude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale) is a prospective cohort of 98,995 women aged 40-65 years at recruitment in 1990 and it constitutes the French component of EPIC. The study was established to produce a large mass of data relevant for the identification of environmental and molecular causes of cancer and other chronic diseases, and to contribute to the development of effective public health strategies. Data on residence, education, early life events, exposures, lifestyle factors and life events were collected at baseline and follow-up questionnaires were sent every 2-3 years with more than 80% response. Mortality data were obtained by means of regular record linkage with various French mortality databases.	Self-reported health outcomes at each questionnaire. All cancers (verified), CVD, diabetes, asthma, osteoporosis, depression, migraine	Metabolomics data on 800 breast cancer cases, 800 controls. B-vitamins 1,000 samples, 1,800 samples, Fatty acids 1,700 samples
NCDS (British 1958 Birth Cohort Study) N= 18,555	The National Child Development Study (NCDS) is a birth cohort established in 1958, which follows the lives of over 17,000 people born in England, Scotland and Wales in a single week in 1958. The participants to the study are	Self-reported health outcomes at each sweep (CVD, Cancer etc). All-cause mortality from register to 2008. Biomedical	Biomarkers from biomedical survey at age 44-46y (lipids, IGE, HbA1c, IGF, CRP, Fibrinogen,

Hypponen et al, 2005 Hypponen et al, 2005	followed up at ages 7, 11, 16, 23, 33, 42, 46 and 50. A biomedical survey (for 9,377 cohort participants) was conducted when participants were aged 44-46 years (N=18,555). The datasets contains prospective SES adult life and early life.	data from measurements (BP; lung function; BMI)	Von Willebrand). Subset: metabolomics data (N=8,000).
Understanding society N=40,000 households (Shiue, I.I, 2014)	<i>Understanding Society</i> is an innovative UK household longitudinal panel study established in 2008 (population-based households sample). This study captures information about the 21 st century UK life, looking specifically at its participants' social and economic circumstances, attitudes, behaviours and health. Information is also collected on all new household members. The data contains information related to SES adult and early life.	Self-reported health data on 20,000 individuals. Biomedical data from nurse visit (6,104).	Blood sample, Saliva sample (details of analytes not released yet)
Gazel (France) N= 20,625 (Westerlund, H., Kivimaki, M., Singh-Manoux, A., Melchior, M., Ferrie JE., Pentti, J., Jokela, M., Leinewebber, C., Goldberg, M., Zins, M. and Vahtera, J.,2009)	The GAZEL study was established in 1989 among employees of the French national gas and electricity company, Electricité de France-Gaz de France (EDF-GDF) (Goldberg et al., 2007, 2015). At baseline (1989), 20,625 employees (15,011 men and 5,614 women), aged 35–50, gave consent to participate. The study design consists of an annual questionnaire used to collect data on health, lifestyle, individual, familial, social and occupational factors, and life events.	Mortality, cancers, CVD events. Diabetes, asthma, depression, sleep disturbances, migraine, quality of life, cognitive and physical functioning. Systematic individual linkage to health insurance, hospital discharge, retirement and mortality national database.	Blood samples (serum, plasma, buffy-coat) for about 5,000 subjects
Constances N (by 2017)= 200,000	The CONSTANCES cohort was established in late 2012 (Zins et al., 2015). It was designed as a randomly selected	Mortality, cancers, CVD events. Diabetes, asthma, depression,	Blood count, glucose, lipids, creatinine, GGT (N=35,000)

May 2016 N: 100,000 Zins, M., Goldberg, M. and Berkman, L., et al, BMC Public Health. 2010	representative sample of French adults aged 18-69 years at inception; 200,000 subjects will be included over a five-year period. At enrolment the participants fill questionnaires collecting data on health, lifestyle, individual, familial, social and occupational factors, and life events and benefits from a comprehensive health examination. The follow-up includes a yearly self-administered questionnaire, a health examination every 5 years and an annual linkage to social and health national databases.	sleep disturbances, migraine, quality of life, cognitive and physical functioning. Systematic individual linkage to health insurance, hospital discharge, retirement (occupational history) and mortality national databases.	
MCCS study (Cancer Council Victoria, Melbourne) N=41,514 Giles, G., Hodge, A. et al, 2013	The Melbourne Collaborative Cohort Study is a prospective cohort study of 41,514 participants (24,469 women) living in Melbourne, Australia. Caucasian volunteers aged between 40 and 69 years were recruited in randomly selected census districts. At baseline (1990–1994), participants attended clinics where demographic, lifestyle and dietary information were collected and anthropometric measurements were performed.	Verified cancer diagnoses, deaths with cause of death, CVD, diabetes, obesity (through physical measures at second visit)	Methylome (Illumina 450K) for 3,000 subjects. Inflammation markers for 500 controls, fatty acids for 1,000 controls. Glucose and lipids for 23,000.
EPITEEN (Portugal) Henrique Barros N=2,942 Ramos, E. et al. ActaPaediatr 2007	Epidemiological Health Investigation of Teenagers in Porto is a random sample of 2,942 adolescents enrolled at public and private schools in Porto, Portugal. The first wave was completed in 2003-2004 when participants were aged 13 years. The second wave in 2007-2008 (17 years), the third wave in 2011-2013 (21 years) and the fourth wave in 2014-2015 (24 years).	Self-reported health outcomes (diseases diagnosis), BP, Anthropometry, Spirometry, Bioelectrical impedance, Bone health (forearm dual-energy X-ray absorptiometry), ECG (21yrs)	CRP, lipids, glucose, insulin

<p>The Cardiovascular Risk in Young Finns Study (Finland) Olli Raitakari N=3,596 (Raitakari, O. et al., 2003)</p>	<p>The Cardiovascular Risk in Young Finns Study (YSF) collected data on 4,320 children and adolescents aged 3, 6, 9, 12, 15 and 18 years at baseline who were randomly chosen from the population register of five Finnish cities with universities with medical schools to produce a representative sample of Finnish children. Of those invited (N=4,320), 3,596 (83%) participated in the first cross-sectional study in 1980. Between 1980 and 1992, these cohorts were followed up at 3-year intervals. The follow-up field studies were performed for the whole study population in 1983 and 1986, when 2,991 (83.2%) and 2,799 (78.3%) subjects participated. In 1989 and 1992 questionnaires were sent to every individual, but only a subset of individuals had their height measured in these survey years so we omit these observations from the analysis. In 2001, 2,283 subjects (63.5%) of the original cohort participated in clinical examinations. A detailed description of the sample can be found elsewhere²². For the purposes of this analysis, we treat YFS as 6 different cohorts representing 3-year age bands (3, 6, 9, 12, 15, and 18 years of age upon recruitment into the study).</p>	<p>Morbidity, mortality, self-reported health outcomes, BP, Anthropometry, Cognitive functioning</p>	<p>CRP, lipids, glucose, insulin, GWAS, metabolomics</p>
---	--	--	--

In addition to the cohorts shown in the table we will also use a large cohort called **Work History Panel (WHIP)**. WHIP is based on a sample of individual-level data from the Social Security Administration archives in Italy, covering almost 8% of all Italian workers employed in the private sector in 1985-

2010. Unlike all other cohorts in **Table 1** it does not have biological samples, but it provides very rich information on income, pensions, unemployment benefits, disability indemnities, workplace and job contracts, linked to hospital and mortality follow-up (see box below).

Cohort	Study design	Outcomes and clinical tests	Already available/funded biomarkers
Work History Panel (WHIP) - Health Italy - N=1,364,922 Costa, G., Leombruni, R. et al, Epidemiol & Prev 2010	The Work History Panel is based on a sample of individual-level data from the Social Security Administration archives in Italy, covering almost 8% of all Italian workers employed in the private sector in 1985-2010. Unlike all other cohorts in LIFEPAH, it does not have biological samples, but it provides very rich information on income, pensions, unemployment benefits, disability indemnities, workplace and job contracts, linked to hospital and mortality follow-up.	Mortality and hospitalization by cause (2001-2012).	Not available.

Figure 2 - Map of the Adult & Child Cohorts participating in LIFEPAATH

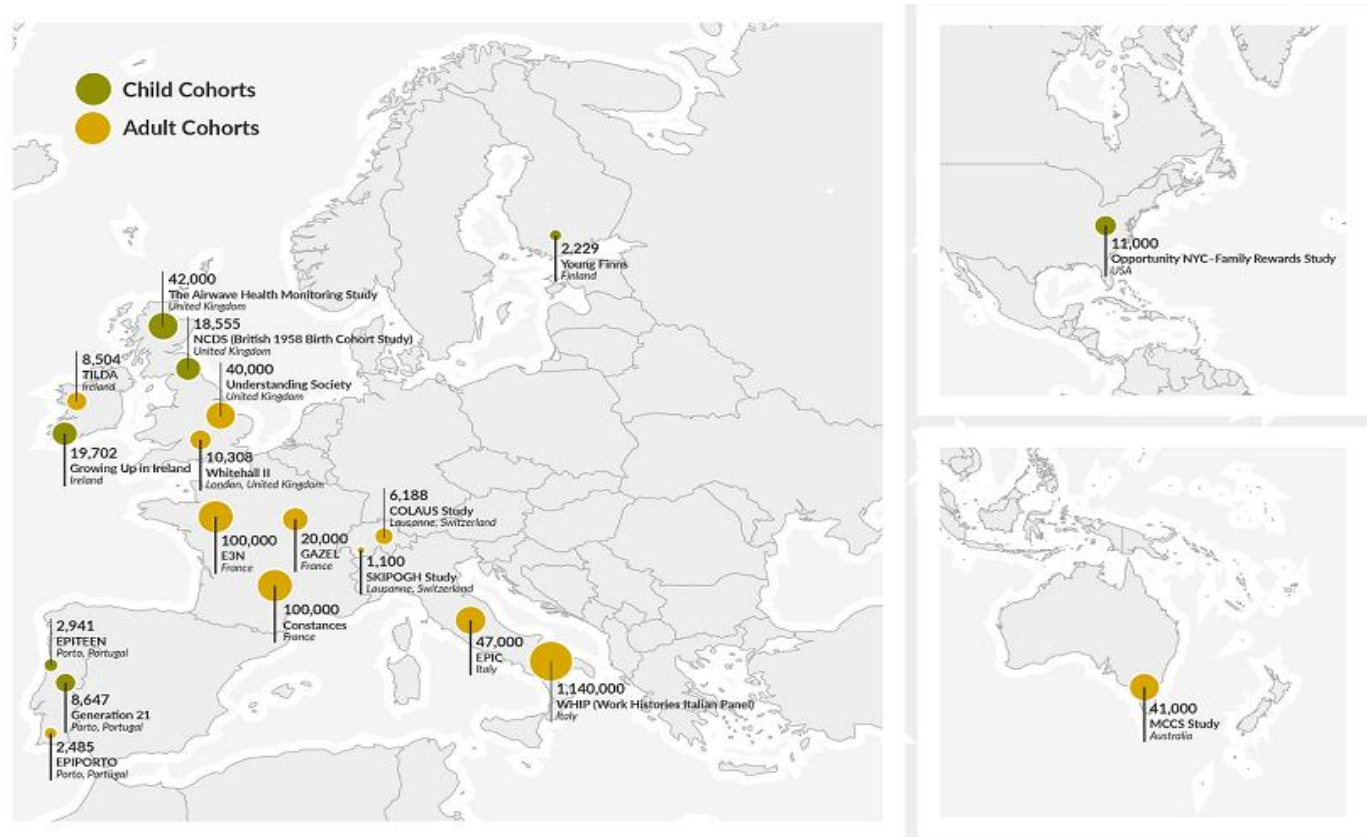


Figure 3. The role of biomarkers in biological pathways leading from SES to healthy ageing.

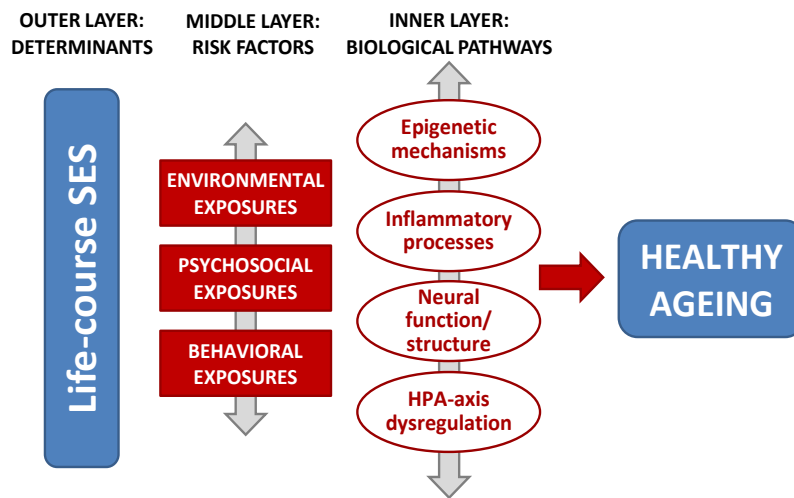
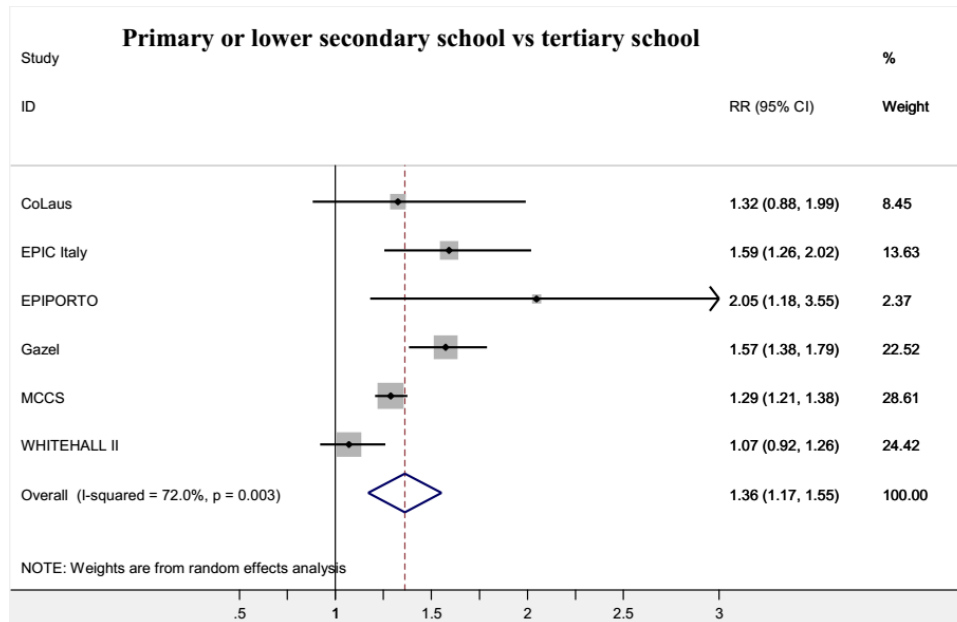


Table 2 – Statistical analyses involved in LIFEPATH

Task	Aim/ Data	Expected outcome relevant to LIFEPATH
SES and exposure variables vs health outcomes	Aim: Harmonize and analyze different indicators of SES available in the cohorts; build a composite indicator of SES; analyse exposure variables. Harmonize and analyze health outcomes and build indicators of healthy ageing	Analysis of all cohorts yielding estimates of: <ul style="list-style-type: none"> - Association between SES measures and healthy ageing variables - Association between SES variables and risk factors - Adjustment of SES-outcome relationships by risk factors
Identification of internal markers of SES	Aim: Identify in the untargeted –omic profiles which candidates are correlated to SES	List of putative internal markers of SES (methodologically validated)
Implementation of Life-course disease risk models	Aim: Identify markers of SES/exposures also relating to the healthy ageing outcomes, and elucidate how their effect is mediated by critical stages in life (early vs late)	Quantified predictive abilities of the validated markers. Estimates of the age-related susceptibility functions, and identification of potential critical age ranges at which SES/exposures are influential with regard to the disease risk.
Burden of disease calculations SES-based risk score	<p>Aim: Estimate the Burden of disease using refined risk estimates based on SES, external exposures and omics.</p> <p>Aim: Develop a prototype for a SES-based score for the prediction of unhealthy ageing.</p>	<p>Updated estimates of the burden of disease for selected exposures.</p> <p>To assess whether the additional predictive ability associated with SES measures can be incorporated into policies and clinical work.</p>

Figure 4 - Meta-analysis of the association of mortality with education, by gender (d'Errico et al, PLOS ONE, *in press*)

MEN



WOMEN

